

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-110186

(43)Date of publication of application : 23.04.1999

(51)Int.Cl.

G06F 3/16

G06F 13/00

G06F 17/30

G06F 17/50

(21)Application number : 09-270161

(71)Applicant : NIPPON TELEGR & TELEPH CORP  
<NTT>

NTT ELECTRONICS CORP

(22)Date of filing : 02.10.1997

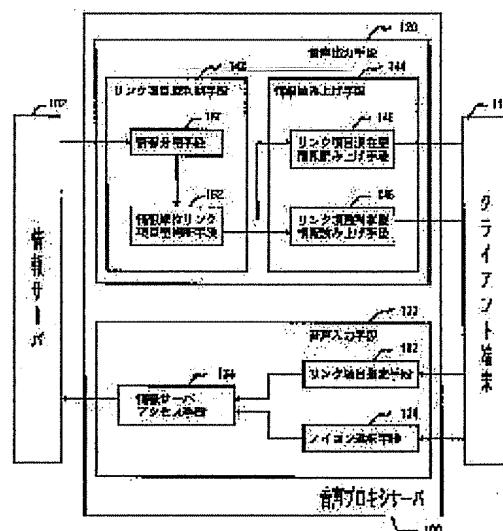
(72)Inventor : HAYASHI HIROMICHI  
KANAMARU TETSUYA  
KIMEDA TSUNEJI  
NAMIKI IKUO  
UJIE MASAMI  
MUSHIKABE KAZUHIKO

(54) BROWSER SYSTEM, VOICE PROXY SERVER, LINK ITEM READING-ALOUD METHOD, AND STORAGE MEDIUM STORING LINK ITEM READING-ALOUD PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a browser system, a proxy server and an information reading-aloud method to express the texts and link items included in the WWW (world wide web) information as audio information.

SOLUTION: This audio proxy server 100 consists of a voice output means 120, which converts the link items into voices to output them to a client terminal 110 via a reading-aloud system according as whether a small number of link items are mixed in the information acquired from an information server 102 (146) or the link items are listed (148) and a voice input means 130, which designates a link item based on the voices inputted from the terminal 110 and accesses the server 102 based on the designated link item.





(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-110186

(43)公開日 平成11年(1999) 4月23日

(51)Int.Cl.<sup>9</sup>  
G 0 6 F 3/16識別記号  
3 4 0F I  
G 0 6 F 3/16

3 4 0 A

3 4 0 N

13/00

3 5 4

13/00

3 5 4 D

17/30

15/403

3 1 0 Z

17/50

15/419

3 2 0

審査請求 未請求 請求項の数21 O L (全 18 頁) 最終頁に続く

(21)出願番号 特願平9-270161

(22)出願日 平成9年(1997)10月2日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(71)出願人 591230295

エヌティティエレクトロニクス株式会社

東京都渋谷区桜丘町20番1号

(72)発明者 林 弘道

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(72)発明者 金丸 哲哉

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(74)代理人 弁理士 伊東 忠彦

最終頁に続く

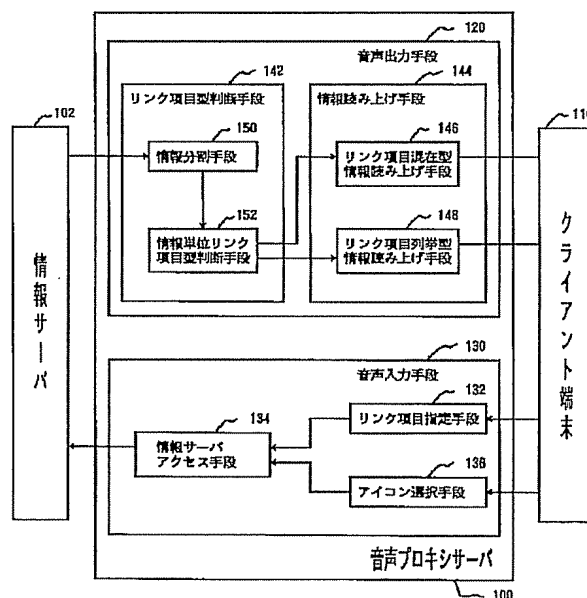
(54)【発明の名称】 ブラウザシステム、音声プロキシサーバ、リンク項目の読み上げ方法及びリンク項目の読み上げプログラムを格納した記憶媒体

(57)【要約】

【課題】 本発明の目的は、WWWの情報に含まれるテキスト及びリンク項目を音声による情報として表現するブラウザシステム、プロキシサーバ及び情報の読み上げ方法を提供することである。

【解決手段】 本発明の音声プロキシサーバ(100)は、情報サーバ(102)から取得された情報に少数のリンク項目が混在する(146)か、又は、リンク項目が列挙されているか(148)に応じた読み上げ方式で、リンク項目を音声に変換してクライアント端末に出力する音声出力手段(120)と、クライアント端末(110)から入力された音声によってリンク項目を指定し(132)、指定されたリンク項目により情報サーバへアクセス(134)する音声入力手段(130)とからなる。

本発明の音声プロキシサーバの構成図



## 【特許請求の範囲】

【請求項 1】 クライアント端末により音声で指定された情報をネットワークを介して情報サーバから取得し、取得された情報を上記クライアント端末に音声で出力するブラウザシステムであって、

上記情報サーバから取得された情報に含まれる関連する情報へのリンク先を示すリンク項目を音声に変換して上記クライアント端末に出力する音声出力手段と、  
上記クライアント端末から入力された音声によってリンク項目を指定し、上記指定されたリンク項目により情報サーバへアクセスする音声入力手段とを有することを特徴とするブラウザシステム。

【請求項 2】 上記音声入力手段は、入力された音声によって上記クライアント端末の画面に表示されたアイコンを選択する手段を更に有することを特徴とする請求項 1 記載のブラウザシステム。

【請求項 3】 上記音声出力手段は、上記情報サーバから取得された情報のテキスト中に含まれるリンク項目の型に応じて、上記情報サーバから取得された情報の型を判断する手段と、  
上記の判断された情報の型に応じて、上記情報を読み上げる手段とを有することを特徴とする請求項 1 又は 2 記載のブラウザシステム。

【請求項 4】 情報の型を判断する手段は、上記情報サーバから取得された情報が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断し、  
上記情報を読み上げる手段は、  
リンク項目混在型の情報を読み上げる手段と、  
リンク項目列挙型の情報を読み上げる手段とを有することを特徴とする請求項 3 記載のブラウザシステム。

【請求項 5】 上記いずれの型であるかを判断する手段は、  
上記情報を所定数以下のリンク項目を含む情報単位に分割する手段と、

上記情報単位が上記リンク項目混在型又は上記リンク項目列挙型のいずれの型であるかを判断する手段とからなることを特徴とする請求項 4 記載のブラウザシステム。

【請求項 6】 クライアント端末により音声で指定された情報をネットワークを介して情報サーバから取得し、取得された情報を上記クライアント端末に音声で出力するブラウザシステムにおいて、

上記情報サーバから取得された情報に含まれる関連する情報へのリンク先を示すリンク項目を音声に変換して上記クライアント端末に出力する音声出力手段と、  
上記クライアント端末から入力された音声によってリンク項目を指定する手段及び上記指定されたリンク項目により情報サーバへアクセスする手段を有する音声入力手段とからなることを特徴とする音声プロキシサーバ。

【請求項 7】 上記音声入力手段は、入力された音声によって上記クライアント端末の画面に表示されたアイコンを選択する手段を更に有することを特徴とする請求項 6 記載の音声プロキシサーバ。

【請求項 8】 上記音声出力手段は、  
上記情報サーバから取得された情報のテキスト中に含まれるリンク項目の型に応じて、上記情報サーバから取得された情報の型を判断する手段と、  
上記判断された情報の型に応じて、上記情報を読み上げる手段とを有することを特徴とする請求項 6 又は 7 記載の音声プロキシサーバ。

【請求項 9】 上記情報の型を判断する手段は、  
上記情報サーバから取得された情報が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断し、  
上記情報を読み上げる手段は、  
リンク項目混在型の情報を読み上げる手段と、  
リンク項目列挙型の情報を読み上げる手段とを有することを特徴とする請求項 8 記載の音声プロキシサーバ。

【請求項 10】 上記いずれの型であるかを判断する手段は、  
上記情報を所定数以下のリンク項目を含む情報単位に分割する手段と、  
上記情報単位が上記リンク項目混在型又は上記リンク項目列挙型のいずれの型であるかを判断する手段とからなることを特徴とする請求項 9 記載の音声プロキシサーバ。

【請求項 11】 情報サーバから関連する情報へのリンク先を示すリンク項目を含むテキスト情報を取得する段階と、  
上記取得されたテキスト情報に含まれるリンク項目が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断する段階と、  
上記判断されたテキスト情報の型に応じて、上記テキスト情報を読み上げる段階とからなることを特徴とするリンク項目を読み上げる方法。

【請求項 12】 上記いずれの型であるかを判断する段階は、  
上記テキスト情報を、所定の数以下の個数のリンク項目が含まれる情報単位に分割し、  
上記分割された各情報単位が、リンク項目混在型又はリンク項目列挙型のいずれの型であるかを判定することを特徴とする請求項 11 記載のリンク項目を読み上げる方法。

【請求項 13】 上記テキスト情報を読み上げる段階は、  
上記情報単位がリンク項目混在型かリンク項目列挙型で

あるかを判別し、

上記情報単位がリンク項目混在型であるならば、リンク項目を含む本文の読み上げと、リンク項目の読み上げとを行い、

上記情報単位がリンク項目列挙型であるならば、リンク項目と対応した本文の読み上げを行うことを特徴とする請求項 1 2 記載のリンク項目を読み上げる方法。

【請求項 1 4】 クライアント端末により音声で指定された情報をネットワークを介して情報サーバから取得し、取得された情報を上記クライアント端末に音声で出力させるプログラムを格納した記憶媒体であって、上記情報サーバから取得された情報に含まれる関連する情報へのリンク先を示すリンク項目を音声に変換させ、上記クライアント端末に音声を出力させる音声出力プロセスと、上記クライアント端末から入力された音声によってリンク項目を指定させ、上記指定されたリンク項目により情報サーバへアクセスさせる音声入力プロセスとを有することを特徴とするリンク項目の読み上げプログラムを格納した記憶媒体。

【請求項 1 5】 上記音声入力プロセスは、入力された音声によって上記クライアント端末の画面に表示されたアイコンを選択させるプロセスを更に有することを特徴とする請求項 1 4 記載のリンク項目の読み上げプログラムを格納した記憶媒体。

【請求項 1 6】 上記音声出力プロセスは、上記情報サーバから取得された情報のテキスト中に含まれるリンク項目の型に応じて、上記情報サーバから取得された情報の型を判断させるプロセスと、上記判断された情報の型に応じて、上記情報を読み上げさせるプロセスとを有することを特徴とする請求項 1 4 又は 1 5 記載のリンク項目の読み上げプログラムを格納した記憶媒体。

【請求項 1 7】 上記情報の型を判断させるプロセスは、上記情報サーバから取得された情報が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断させ、上記情報を読み上げるプロセスは、リンク項目混在型の情報を読み上げさせるプロセスと、リンク項目列挙型の情報を読み上げさせるプロセスとを有することを特徴とする請求項 1 6 記載のリンク項目の読み上げプログラムを格納した記憶媒体。

【請求項 1 8】 上記いずれの型であるかを判断させるプロセスは、上記情報を所定数以下のリンク項目を含む情報単位に分割させるプロセスと、上記情報単位が上記リンク項目混在型又は上記リンク項目列挙型のいずれの型であるかを判断させるプロセスと

を有することを特徴とする請求項 1 7 記載のリンク項目の読み上げプログラムを格納した記憶媒体。

【請求項 1 9】 情報サーバから関連する情報へのリンク先を示すリンク項目を含むテキスト情報を取得させるプロセスと、上記取得されたテキスト情報に含まれるリンク項目が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断させるプロセスと、

上記判断されたテキスト情報の型に応じて、上記テキスト情報を読み上げさせるプロセスとからなることを特徴とするリンク項目を読み上げるプログラムを格納した記憶媒体。

【請求項 2 0】 上記いずれの型であるかを判断させるプロセスは、上記テキスト情報を、所定の数以下の個数のリンク項目が含まれる情報単位に分割させ、上記分割された各情報単位が、リンク項目混在型又はリンク項目列挙型のいずれの型であるかを判定させることを特徴とする請求項 1 9 記載のリンク項目を読み上げるプログラムを格納した記憶媒体。

【請求項 2 1】 上記テキスト情報を読み上げさせるプロセスは、上記情報単位がリンク項目混在型かリンク項目列挙型であるかを判別させ、上記情報単位がリンク項目混在型であるならば、リンク項目を含む本文の読み上げと、リンク項目の読み上げとを行わせ、上記情報単位がリンク項目列挙型であるならば、リンク項目と対応した本文の読み上げを行わせることを特徴とする請求項 2 0 記載のリンク項目を読み上げるプログラムを格納した記憶媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】本発明は、コンピュータとネットワークからなるクライアント／サーバ構成の情報提供システム、とりわけインターネットにおけるWWW(World Wide Web)システムにおいて、取得したい情報を指定する指定情報をクライアント端末のマイクから音声で入力し、サーバに蓄積されてる情報を取得し、クライアント端末に音声で出力する音声ブラウザシステムに関する。

【0 0 0 2】

【従来の技術】周知のように、WWWシステムとして、サーバ及びクライアントのハードウェア、ソフトウェアがネットワーク上に適切に構成されている場合に、クライアント端末上にインストールされた例えば、Netscape Navigator™等のブラウザを使用することによって、サーバに格納したテキストやイメージの情報をクライアン

ト画面上に表示して閲覧することが可能である。このようなシステムの場合に、画面上の特定の情報をマウスなどで選択すると、この特定の情報と関連づけられた情報にアクセスし、画面上に表示して閲覧することが可能である。以下では、特定の情報と「関連づけられる」ことを「リンクを張られた」と称し、上記「特定の情報」を「リンク項目」と称する。このようなシステムでは、情報が視覚的情報として提供されることを前提とする。

【0003】従って、情報のサービスを享受するためには画面に目を向ける必要があるので、視覚の不自由な人は情報のサービスを享受することができないという問題がある。

【0004】

【発明が解決しようとする課題】本発明は第一に、視覚の不自由な人に対しても上記の情報サービスの利用を可能にすることを目的とする。即ち、最近の音声認識技術及び音声合成技術を使用することによって、リンク項目の指定等をマイクからの音声により行い、サーバから取得した情報を、画面上に表示するのではなく、音声合成音で出力することで、視覚の不自由な人の利用に供することを可能にする。具体的には、ブラウザのアイコン

(例えば、前進、後退等)の選択及びリンク項目の指定が音声入力により行なわれ、入力された音声情報が認識され、認識された音声情報により指定されたURL(Uniform Resource Locator)がアクセスされ、情報が取得され、取得された情報の中のテキスト部分がクライアント端末のスピーカから合成音として出力される。

【0005】本発明は第二に、視覚の不自由な人に情報サービスを提供する場合、リンク項目の出力及び入力を音声で行う場合の使い勝手を良くすることを目的とする。即ち、WWWの情報には、長い文章が含まれている場合、或いは、10乃至20個のリンク箇所が随所に指定されている場合がある。また、実際には、カラーのイメージ情報がテキストに混在するだけでなく、動画とリンクが張られ、視覚に訴える情報がふんだんに使用されている。従って、通常は画面に表示されるWWWの情報を聴覚的な情報として表現する場合に、単にテキスト部分だけを読み上げても、取得された情報の内容を理解することが難しく、その結果として、利用者が次に取得したい情報を指定するためのリンク項目の選択が難しくなるという問題がある。そのため、視覚の不自由な人はWWWの情報のサービスを十分に享受できない。

【0006】本発明は、上記従来技術の問題点に鑑み、WWWの情報に含まれるテキスト及びリンク項目を音声による情報として表現する改良されたブラウザシステム、プロキシサーバ及び情報の読み上げ方法の提供を目的とする。

【0007】

【課題を解決するための手段】図1は本発明の原理構成図である。本発明では、上記の課題を解決するために、

同図に示す如く、クライアント端末110により音声で指定された情報をネットワーク104を介して情報サーバ102から取得し、取得された情報を上記クライアント端末110に音声で出力するブラウザシステムは、情報サーバ102から取得された情報に含まれる関連する情報へのリンク先を示すリンク項目を音声に変換してクライアント端末110に出力する音声出力手段120と、クライアント端末110から入力された音声によってリンク項目を指定し、指定されたリンク項目により情報サーバ102へアクセスする音声入力手段130とを有することを特徴とする。

【0008】上記音声入力手段120により、音声による入力からリンク項目及びアイコンを指定することが可能になり、かつ、上記音声出力手段130により、クライアント端末に表示される情報を読み上げることが可能になる。図2は本発明の音声プロキシ(Proxy)サーバ100を表わす図である。本発明は、情報サーバ102から取得された情報に含まれる関連する情報へのリンク先を示すリンク項目を音声に変換してクライアント端末110に出力する音声出力手段120と、クライアント端末110から入力された音声によってリンク項目を指定する手段132及び指定されたリンク項目により情報サーバ102へアクセスする手段134を有する音声入力手段130とからなることを特徴とする音声プロキシサーバ100である。

【0009】上記音声プロキシサーバ100の音声入力手段130は、入力された音声によって上記クライアント端末の画面に表示されたアイコンを選択する手段136を更に有する方が有利である。上記音声出力手段120は、上記情報サーバから取得された情報のテキスト中に含まれるリンク項目の型に応じて、上記情報サーバから取得された情報の型を判断する手段142と、上記判断された情報の型に応じて、上記情報を読み上げる手段144とを有する。

【0010】上記情報の型を判断する手段142は、上記情報サーバから取得された情報が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断する。この判断に対応して、上記情報を読み上げる手段144は、リンク項目混在型の情報を読み上げる手段146と、リンク項目列挙型の情報を読み上げる手段148とを有する。

【0011】上記いずれの型であるかを判断する手段142は、上記情報を所定数以下のリンク項目を含む情報単位に分割する手段150と、上記情報単位が上記リンク項目混在型又は上記リンク項目列挙型のいずれの型であるかを判定する手段152とからなる点が有利である。本発明の音声プロキシサーバ100により、音声による入力からリンク項目及びアイコンを指定することが可能になり、かつ、クライアント端末に表示される情報

を読み上げることが可能になる。

【0012】更に、本発明によれば、ブラウザ上に表示したインターネット情報のアイコン、テキスト本文、リンク項目の読み上げ方が、テキスト本文中のリンク項目の数及びリンク項目の型（リンク項目混在型又はリンク項目列挙型）に従って判断され、インターネット情報が自動的に処理されて音声出力される点に特徴がある。これにより、視覚的な情報を用いることなく、リンク項目を簡単に選択できるようになる。

【0013】図3は本発明によるリンク項目を読み上げる方法の原理を説明する図である。本発明のリンク項目を読み上げる方法は、情報サーバから関連する情報へのリンク先を示すリンク項目を含むテキスト情報を取得する段階（ステップ100）と、上記取得されたテキスト情報に含まれるリンク項目が、テキスト中に少数のリンク項目が含まれるリンク項目混在型、又は、テキスト中にリンク項目が列挙されているリンク項目列挙型のいずれの型であるかを判断する段階（ステップ110）と、上記判断されたテキスト情報の型に応じて、上記テキスト情報を読み上げる段階（ステップ120）とからなることを特徴とするリンク項目を読み上げる方法。

【0014】上記いずれの型であるかを判断する段階（ステップ110）は、上記テキスト情報を、所定の数以下の個数のリンク項目が含まれる情報単位に分割し（ステップ111）、上記分割された各情報単位が、リンク項目混在型又はリンク項目列挙型のいずれの型であるかを判断する（ステップ112）方が有利である。また、上記テキスト情報を読み上げる段階（ステップ120）は、上記情報単位がリンク項目混在型かリンク項目列挙型であるかを判別し（ステップ121）、上記情報単位がリンク項目混在型であるならば、リンク項目を含む本文の読み上げと、リンク項目の読み上げとを行い（ステップ122）、上記情報単位がリンク項目列挙型であるならば、リンク項目と対応した本文の読み上げを行う（ステップ123）ことを特徴とする。

【0015】

【発明の実施の形態】図4は本発明の第1の実施例の音声ブラウザシステムの構成図である。同図に示された音声ブラウザシステムのクライアント端末は、クライアント端末本体1と、クライアント端末本体1に接続された音声入力用のマイク2及び音声合成音などの音声出力用のスピーカ3と、クライアント端末本体1に接続された表示装置4とからなる。情報サーバであるWWWサーバ7から取り出された情報は、通常、表示装置4に表示される。表示装置4は、視覚の不自由な人には直接有効ではない。

【0016】音声ブラウザシステムは、音声ブラウザのための処理を行う音声プロキシサーバ5を更に有する。クライアント端末はネットワーク6を介して音声プロキシサーバ5に接続される。ネットワーク6は、周知のイ

ーサネット、ATM等のいずれのネットワークでもよく、本発明はネットワークの形態に限定されるものではない。情報を提供するWWWサーバ7は、世界中に接続されているインターネット8に接続されている。

【0017】図3は、本発明の第1の実施例のサーバクライアント構成を示す図である。クライアント21は、例えば、Netscape Navigator<sup>TM</sup>のようなブラウザ22、ブラウザ22との間で音声の入出力を行う音声入出力AP（アプリケーションプログラム）23、音声入出力AP23からブラウザに自動的に任意の情報先（通常、URLと称される）へアクセスさせるための制御を行うブラウザ制御アプリケーションプログラム24、及び、音声入力の際に、ボタンの入力を監視するボタン監視アプリケーションプログラム25からなる。但し、音声のパワーを検出することにより上記ボタンを使用しないで音声入力を監視しても構わない。

【0018】音声プロキシサーバ26は、マイクからの音声を認識する音声認識エンジン28と、テキストを音声に合成する音声合成エンジン29と、文章を名詞又は複合語単位に分解する形態素解析エンジン33と、クライアント21とWWWサーバ32との中継、並びに、音声認識エンジン28及び音声合成エンジンとのインタフェース処理とを行う音声プロキシ部27とからなる。音声プロキシ部27は、音声出力、即ち、読み上げ条件に基づく読み上げを行う読み上げ前処理部30を有する。

【0019】本発明の第1の実施例によれば、漢字コード変換機能やキャッシュメモリ機能を有する中継サーバ31が設けられているが、これらの機能は音声プロキシサーバ26の中に組み込んでもよい。WWWサーバ32は既存のhttpdである。図6は図5に示した本発明の第1の実施例のサーバクライアントの動作シーケンスチャートである。以下の説明では、利用者がWWWサーバに入り、最初のページをアクセスし、その後に任意のページにユーザがアクセスする場合を考える。各動作について順次説明する。

【0020】ステップ1）利用者がクライアント本体、例えば、パソコンの電源を入れると、予め電源投入時に起動するよう設定された音声入出力アプリケーションプログラム23が自動的に起動される。このような起動は、パソコンの設定方法により実現される。

ステップ2）音声入出力アプリケーションプログラム23は、Netscape Navigator<sup>TM</sup>のようなブラウザ22を起動させる。

【0021】ステップ3）ブラウザ22は起動時に表示すべきURLを取得する要求を音声プロキシ部27に送る。

ステップ4）音声プロキシ部27はブラウザ22の要求するURLの要求をWWWサーバ32に発行する。

ステップ5）WWWサーバ32のhttpdは要求されたURLのデータを返す。

【0022】ステップ6) 音声プロキシ部27はURLをブラウザ23に返し、ブラウザ23はURLをパソコン画面に表示する。

ステップ7) 音声プロキシ部27の前処理部30は、URLのデータに基づいて読み上げ前処理を行う。読み上げ前処理とは、HTMLデータから、クライアントで音声出力させるテキストを抽出し、プレーンテキスト化し、タイトル部、本文、リンク項目等に分類すると共に、ガイダンスを追加して音声合成エンジン29に音声変換を要求する処理である。但し、起動時に表示されるURLの場合には、本文及びリンク案内ガイダンスの音声変換だけが要求される。

【0023】ステップ8) 音声合成を依頼された音声合成エンジン29は、テキストデータを音声データに変換し、音声プロキシ部27に渡す。

ステップ9) 音声プロキシ部27は、音声合成エンジン29から受けた音声データをクライアントの音声入出力アプリケーションプログラム23に伝達し、音声入出力アプリケーションプログラム23は音声データを音声出力する。

【0024】ステップ10) 音声プロキシ部27は、読み上げ前処理で抽出されたリンク項目の形態素解析を形態素解析エンジン33に要求する。

ステップ11) 形態素解析エンジン33は、リンク項目の1文章を名詞単位に分解して、音声認識時の候補として抽出すると共に、名詞が連続する際には、複合語も候補として抽出し、解析結果を音声プロキシ部27に渡す。形態素解析の技術は、発明の名称「形態素解析方法および装置」の特願平7-291143号明細書又は発明の名称「ハイパーテキスト中継方法及び装置」の特願平8-312016号明細書に記載されている。例えば、「内閣総理大臣が渡米」というリンク項目がある場合に、このステップの形態素解析により得られる音声認識時の候補は、全文の「内閣総理大臣が渡米」、「内閣」、「大臣」、「内閣総理」、「総理大臣」、「内閣総理大臣」及び「渡米」である。

【0025】ステップ12) 音声プロキシ部27は、形態素解析エンジン33の解析結果を音声認識エンジン28に渡し、辞書登録を行う。

ステップ13) 次に、利用者は音声入力を開始するため、ボタン(図示しない)を押下する。クライアント21のボタン監視アプリケーションプログラム25は、ボタンが押下されたので、音声入出力アプリケーションプログラム23に通知する。このタイミングで音声入出力アプリケーションプログラム23は音声の録音を開始する。

【0026】ステップ14) 更に、ボタン監視アプリケーションプログラム25は、ボタンを監視し、ボタンが放されたことを検出したタイミングで、音声入出力アプリケーションプログラム23に通知し、音声入出力アプ

リケーションプログラム23は音声の録音を終了し、音声入出力アプリケーションプログラム23へ通知する。

【0027】ステップ15) 音声入出力アプリケーションプログラム23は入力された音声を音声プロキシ部27に送信する。

ステップ16) 音声プロキシ部27は、受信した音声データに基づいて音声認識エンジン28に音声認識を要求する。

ステップ17) 音声認識エンジン28は、音声データを受信し、音声認識を実行して結果を音声プロキシ部27に返す。

【0028】ステップ18) 音声プロキシ部27は、認識結果を基にして、現在表示されている情報から要求されたリンク項目のURLを取得し、WWWサーバ32へ要求を発行する。

ステップ19) WWWサーバ32のhttpdは、要求されたURLのデータを返す。

【0029】ステップ20) 音声プロキシ部27の前処理部30は、WWWサーバ32から返されたURLのデータに基づいて読み上げ前処理を行う。ここで、ブラウジング案内ガイダンスだけの音声合成を音声合成エンジン29に要求する。

ステップ21) 音声合成エンジン29は、ブラウジング案内ガイダンスを音声データに変換し、音声プロキシ部27に渡す。

【0030】ステップ22) 音声プロキシ部27は、受信した音声データをクライアント21の音声入出力アプリケーションプログラム23に送り音声出力を要求すると同時に、インターネット情報の表示を行う命令を発行する。

ステップ23) 音声入力アプリケーションプログラム23は、音声出力を行うと同時に、ブラウザ制御アプリケーションプログラム24へインターネット情報の表示を行う命令を発行する。

【0031】ステップ24) ブラウザ制御アプリケーションプログラム24は、音声入力アプリケーションプログラム23の要求通りにブラウザにインターネット情報の表示を行う命令を発行する。

ステップ25) ブラウザ22は、指定されたURLを取得する要求を音声プロキシ部27に出す。

【0032】ステップ26) 音声プロキシ部27は、そのブラウザ22の要求に対し、ステップ19)で受信されたURLデータを速やかに返送し、ブラウザ22はそのデータを表示する。

ステップ27) 音声プロキシ部27の前処理部30は、URLのデータに基づいて読み上げ前処理を行う。ここでは、本文及びリンク項目の音声変換だけを音声合成エンジン29に要求する。

【0033】ステップ28) 音声合成エンジン29は、テキストデータを音声に変換し、音声プロキシ部27に



渡す。

ステップ 29) 音声プロキシ部 27 は受信した音声データを音声入出力アプリケーションプログラム 12 に渡し、音声入出力アプリケーションプログラム 12 は音声出力を行う。

【0034】ステップ 30) 音声プロキシ部 27 は、読み上げ前処理部 30 で抽出されたリンク項目の形態素解析を形態素解析エンジン 33 に要求する。

ステップ 31) 形態素解析エンジン 33 は、ステップ 11) と同様に、リンク項目の形態素解析を行い、解析結果を音声プロキシ部 27 に渡す。

ステップ 32) 音声プロキシ部 27 は、形態素解析エンジン 33 の解析結果を音声認識エンジン 28 に渡し、辞書登録を行う。

【0035】これにより、サーバクライアントシステムは、利用者からの音声入力待ちになり、ステップ 13) 以降の処理が繰り返行われる。

【0036】

【実施例】図 7 は、本発明の第 2 の実施例において、WWWサーバから取り出され、ブラウザ上に表示された情報の型の例を表わす図である。同図の (A) は、テキスト中に少数のリンク項目が含まれるリンク項目混在型の場合を表わす。タイトル欄 41 に示された「音声ブラウザ」がこの情報のタイトルである。本文欄 42 には、この情報の本文が示される。本文の一部には、リンクが張られたリンク項目 43 が示される。この例では、影文字で示された「ブラウザ」、「表示装置」及び「インターネット」の 3 箇所にリンクが張られている。これらのリンク項目を選択することにより、夫々のリンクが張られた先の情報にアクセスできる。本文欄 42 内の枠に示された画像は、この情報のイメージ部分 44 である。

【0037】図 7 の (B) にはリンク項目列挙型の場合の情報の例が示される。この例の場合、多数のリンク項目 45 が列挙されている。この例は、WWWサーバから取り出された情報が検索結果の一覧の場合であり、情報の中の殆どの部分がリンク項目の列挙に該当している。例えば、目次は、このように殆どの部分がリンク項目の列挙である。

【0038】情報を表示形式で分類すると、図 7 に示された混在型と列挙型の 2 種類がある。取り出された情報を音声で読み上げる方式は、情報の型によって異なる。図 8 は本発明の第 2 の実施例によるリンク項目混在型の情報の読み上げ順方式を表わす図である。同図に示された例は、図 7 の (A) に示された情報に基づいている。

【0039】図 8 の (A) は、タイトル、リンク項目、本文の順に読み上げるタイトルーリンクー本文順読み上げ方式を表わす図である。この方式の場合、最初に、情報のタイトル「音声ブラウザ」の読み上げ部分 51 が読み上げられる。情報のタイトルの読み上げ部分 51 には、アクセスした情報のタイトルが挿入される。「ア

セスします」の部分 52 は、タイトル「音声ブラウザ」にアクセスすることをガイダンスしているところである。この方式では、この部分 52 は、例えば、プログラムで固定的に使用される。尚、図面及び明細書の説明中、下線付きの部分は読み上げ時に固定的に使用される箇所を表わす。読み上げ部分 51 と、固定部分 52 とから、「音声ブラウザにアクセスします」と読み上げられる。ここで、固定部分 52 の「アクセスします」は、例えば、「という情報を取り出し表示します」のように置き換えても構わない。このようにガイダンス文自体は、自由な形にしてよいことはいふまでもない。次に、リンク項目読み上げのガイダンス文 53 が続く。更に、この情報の中でリンクが張られている全項目、例えば、「インターネット」の項目 55 等の読み上げ部分 54 である。この例では、「ブラウザ」、「表示装置」及び「インターネット」の 3 項目が含まれる。全項目の読み上げ部分 54 に続いて、本文読み上げ開始のガイダンス部分 56 と、本文自体の読み上げ部分 57 とがある。

【0040】図 8 の (B) は、タイトルー本文ーリンク順読み上げ方式を表わす図である。同図の (A) のタイトルーリンクー本文順読み上げ方式とは異なり、本文を先に読み上げた後に、リンク項目が読み上げられる。同図の (C) は、タイトルー本文ーリンク順読み上げ方式において、リンク項目が自動的に読み上げられない方式に相当する。即ち、本文の読み上げ終了の時点で、音声によって「リンク項目」と入力することにより、その情報のリンク項目が読み上げられる方式である。

【0041】一方、列挙型の情報の場合には、リンク項目と本文とが略一致するため、上記の混在型情報の読み上げ順方式をそのまま使用することはできない。図 9 は、本発明の第 2 の実施例による列挙型情報の読み上げ順方式を表わす図である。図 8 に示された混在型と比較すると、リンク項目が本文と一致した場合に相当する。図 9 には、列挙型の読み上げ順方式に固有のガイダンス文 58 が示される。それ以外は図 8 に示された混在型と同様に読み上げられる。

【0042】図 10 は、本発明の第 3 の実施例による読み上げ条件自動処理のフローチャートである。情報が WWWサーバから取り出されると、音声プロキシサーバ上の蓄積装置に情報が一時的に格納される。蓄積装置に蓄積された情報を必要に応じてメモリに転送し、読み上げ条件自動処理を開始する (ステップ 51)。

【0043】最初に、与えられた読み上げ条件を参照して、読み上げの際に分割される情報単位に含まれるリンク項目のリンク数  $L_n$  と、読み上げ条件と、情報型判断指数  $A$  とを設定する (ステップ 52)。ここで、読み上げ条件とは、リンク項目又は本文のいずれを先に読み上げるかを指定する条件である。また、情報型判断指数とは、分割された本文 (以下、分割本文と称する) 内のリンク項目の総文字数と分割本文の総文字数との基準の比

を表わす量である。

【0044】次に、分割本文内のリンク項目数が取得された分割のリンク数 $L_n$ 以下になる場所、かつ、文章単位で本文を先頭から分割する(ステップ53)。本文が分割された後、分割本文がリンク項目列挙型或いはリンク項目混在型のいずれであるかが判定される(ステップ54)。判定の方法は、分割本文内のリンク項目の総文字数( $p$ )と分割本文の総文字数( $q$ )との比( $p/q$ )と、情報型判断指数( $A$ )との大小関係による。即ち、 $p/q > A$ ならば、列挙型であると判定され、 $p/q < A$ であるならば、混在型であると判定される。通常、情報型判断指数 $A$ は、限りなく1に近い値が設定される。

【0045】分割本文の情報が混在型であると判定された場合、次に、リンク項目と本文のいずれを先に読み上げるかを読み上げ条件に基づいて判定する(ステップ55)。リンク項目が先に判定されたならば、 $L_n$ 個のリンク項目が読み上げられ(ステップ56)、続いて、分割本文が読み上げられる(ステップ57)。

【0046】一方、本文が先に判定されたならば、まず分割本文が読み上げられ(ステップ58)、次に、 $L_n$ 個のリンク項目が読み上げられる(ステップ59)。分割本文の情報が列挙型であると判定された場合、 $L_n$ 個分のリンク項目と、分割本文とが一度に読み上げられる(ステップ60)。リンク項目と分割本文とが読み上げられた後、本文がすべて終了したかどうか判定される(ステップ61)。

【0047】未だ読み上げられていない分割本文がある場合、次の分割本文が取り出され(ステップ62)、ステップ54に戻り、次の分割本文について同様に読み上げ条件自動処理が繰り返される。本文がすべて終了している場合には、読み上げ終了の処理に進む(ステップ63)。図11は本発明の第4の実施例による分割本文とその読み上げ内容とを示す図である。本文の分割と、読み上げ内容の決定は、音声プロキシサーバの音声プロキシ部に設けられた読み上げ前処理部で行われる。

【0048】読み上げ条件の内容61は、本文の分割がリンク数=3、読み上げ順序がタイトル、リンク項目、本文の順であることを示している。情報型判断指数 $A=0.95$ である。設定内容を設定する方法は、例えば、画面から設定されたデータをファイルに書き込む方法、或いは、ファイルに直接書き込む方法等のいずれの方法でも良く、本発明は読み上げ条件の内容の設定方法によって限定されるものではない。

【0049】情報の本文全体62は、リンク数=3を用いて、分割本文1、分割本文2及び分割本文3の3つに分割されていることが分かる。更に、分割本文の中で、分割本文1及び分割本文3はリンク項目混在型であり、分割本文2はリンク項目列挙型である。読み上げ内容63は、このような分割本文に従って生成された読み上げ

例を表わす。

【0050】混在型の分割本文1及び分割本文3は、読み上げ条件に従って、リンク項目、本文の順に読み上げられることが分かる。列挙型の分割本文3の場合には、リンク項目と本文とが同一であるため、まとめて一度だけ読み上げられる。また、次の分割本文に移る場合には、音声コマンド「次ぎ」を使用してもよい。また、音声プロキシサーバ26の構成は、上記の実施例で説明された例に限定されることなく、音声プロキシサーバ26の各々の構成要件をソフトウェア(プログラム)で構築し、ディスク装置等に格納しておき、必要に応じて情報提供装置のコンピュータにインストールしてリンク項目の読み上げを行うことも可能である。さらに、構築されたプログラムをフロッピーディスクやCD-ROM等の可搬記憶媒体に格納し、このようなシステムを用いる場面で汎用的に使用することも可能である。

【0051】本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【0052】

【発明の効果】以上に詳述したように、本発明によれば、ブラウザ上に取り出された情報のうちのテキスト本文及びリンク項目の読み上げ方が、利用者の設定内容であるリンク項目の数と、リンク項目が情報に混在的に含まれるか、又は、列挙的に含まれるかを表わす情報の型とに応じて、自動的に処理され、テキスト本文及びリンク項目が音声出力されるので、テキスト本文及びリンク項目の読み上げ方が利用者の好みの条件で使用されるという利点がある。

【0053】本発明の音声プロキシサーバにより、音声による入力からリンク項目及びアイコンを指定することが可能になり、かつ、クライアント端末に表示される情報を読み上げることが可能になる。従って、通常のブラウザを持つクライアント端末は、音声認識或いは音声合成のような特別なソフトウェアを別途準備することなく、音声を利用したブラウザシステムを実現することができる。

【0054】更に、クライアント端末ではなく、音声プロキシサーバに音声認識及び音声合成の手段を設けることにより、クライアント端末の機種に殆ど依存することのない汎用的なブラウザシステムを構築することが可能になる。また、プロキシサーバは大規模な辞書を搭載することが可能であり、クライアント端末に記事を表示している間に先行して各種変換処理を行うことが可能である。従って、テキスト系サービスのサーバ、エンジンを拡張する際に、例えば、翻訳、要約のような各種変換処理されたテキストを同時に配信することが可能である。

【図面の簡単な説明】

【図1】本発明の原理構成図である。

【図2】本発明の音声プロキシサーバの構成図である。

15

【図 3】本発明のリンク項目の読み上げ方法の原理説明図である。

【図 4】本発明の第 1 の実施例による音声ブラウザシステムの構成図である。

【図 5】本発明の第 1 の実施例のサーバクライアントシステムの構成図である。

【図 6】本発明の第 1 の実施例のサーバクライアントシステムの動作シーケンスチャートである。

【図 7】本発明の第 2 の実施例における情報の型の例の説明図である。

【図 8】本発明の第 2 の実施例における混在型情報の読み上げ順の説明図である。

【図 9】本発明の第 2 の実施例における列挙型情報の読み上げ順の説明図である。

【図 10】本発明の第 3 の実施例による読み上げ条件自動処理のフローチャートである。

【図 11】本発明の第 4 の実施例による分割本文と読み

16

上げ内容との説明図である。

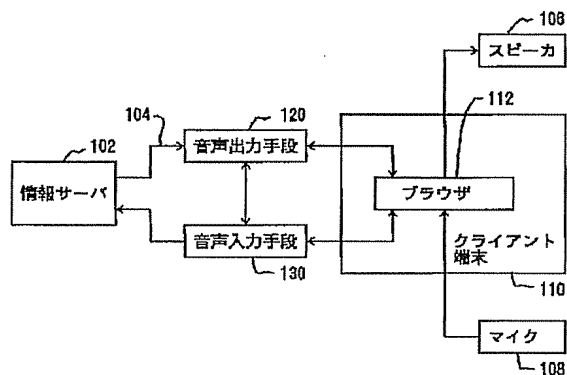
【符号の説明】

- 100 音声プロキシサーバ
- 102 情報サーバ
- 110 クライアント端末
- 120 音声出力手段
- 130 音声入力手段
- 132 リンク項目指定手段
- 134 情報サーバアクセス手段
- 136 アイコン選択手段
- 142 リンク項目型判断手段
- 144 情報読み上げ手段
- 146 リンク項目混在型情報読み上げ手段
- 148 リンク項目列挙型情報読み上げ手段
- 150 情報分割手段
- 152 情報単位リンク項目型判断手段

10

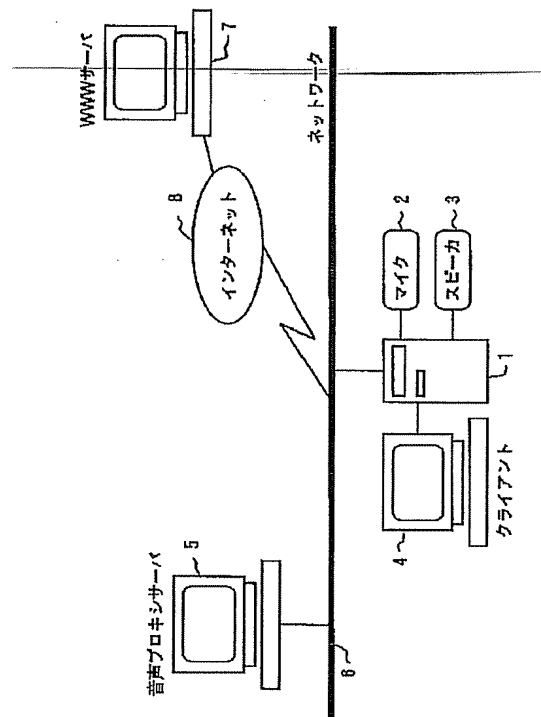
【図 1】

本発明の原理構成図



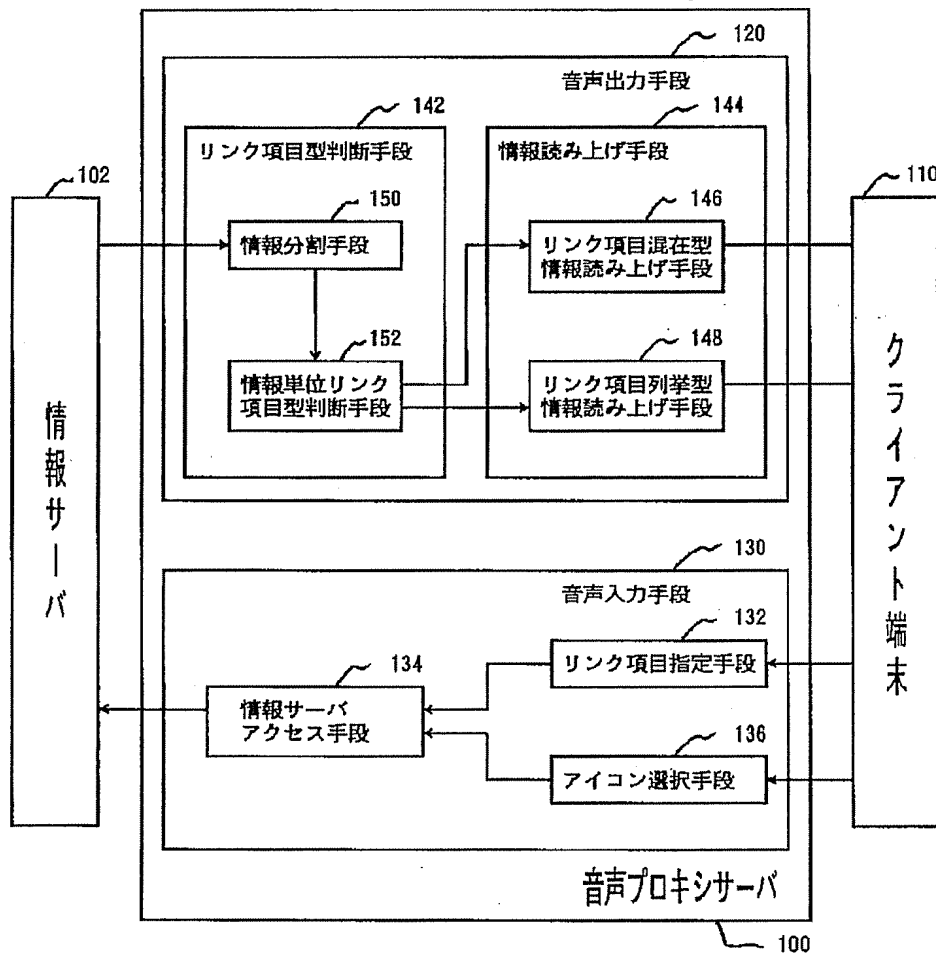
【図 4】

本発明の第 1 の実施例による音声ブラウザシステムの構成図



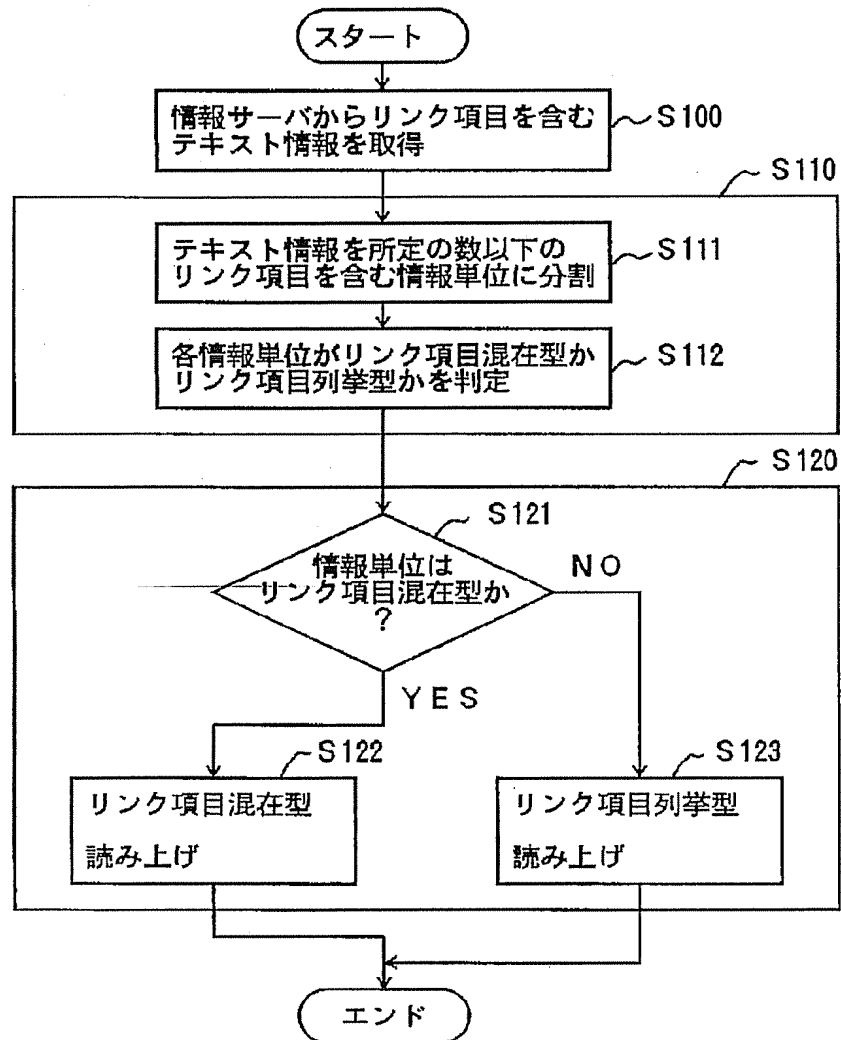
【図2】

## 本発明の音声プロキシサーバの構成図



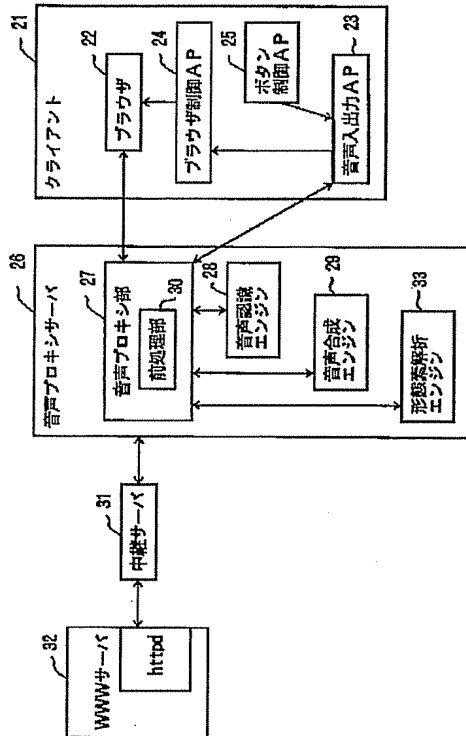
【図3】

## 本発明の原理説明図

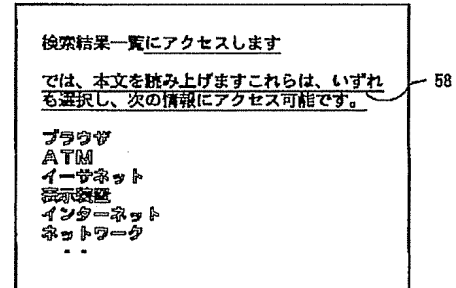


【図5】

本発明の第1の実施例のサーバクライアントシステムの構成図



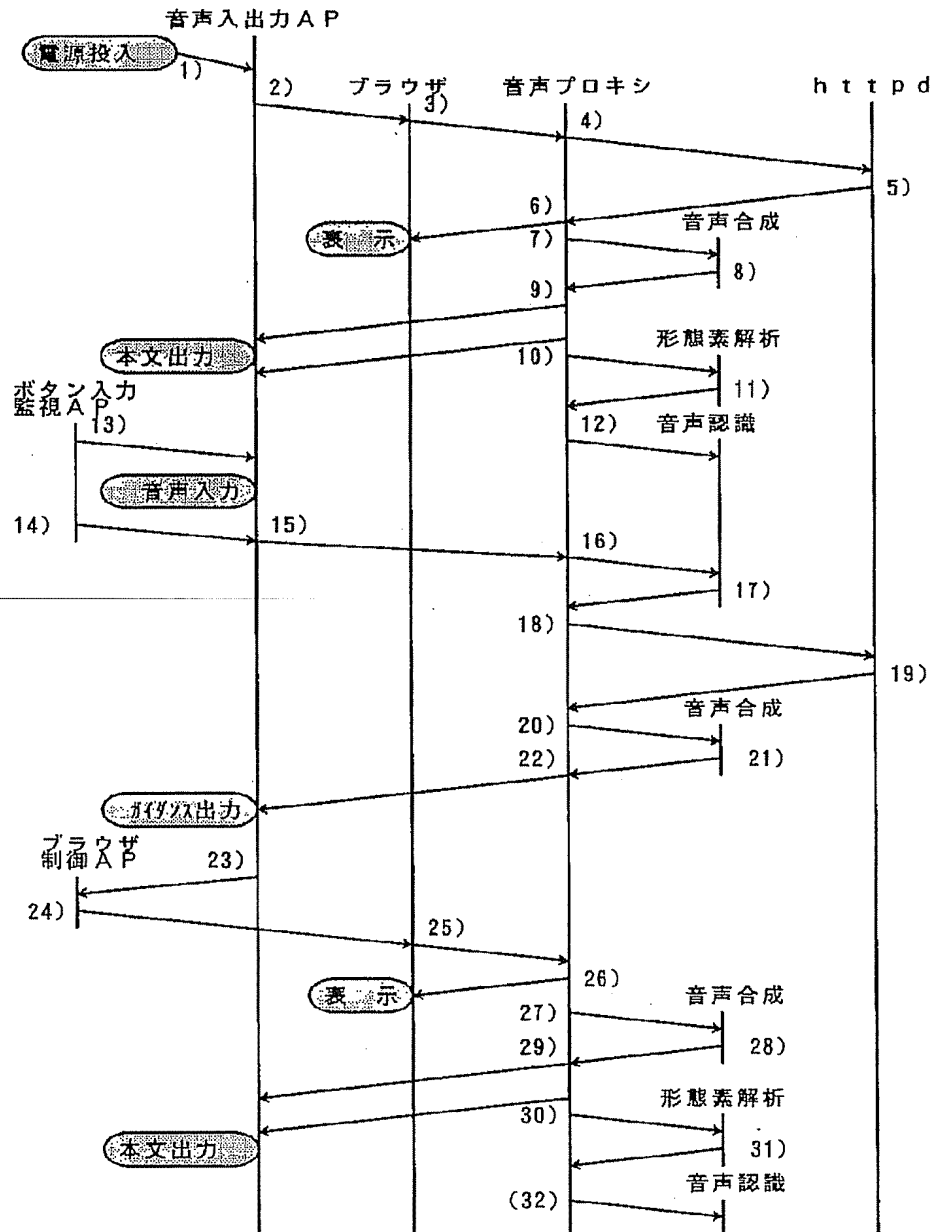
【図9】

本発明の第2の実施例における  
列挙型情報の読み上げ順の説明図

タイトルリンク順読み上げ方式

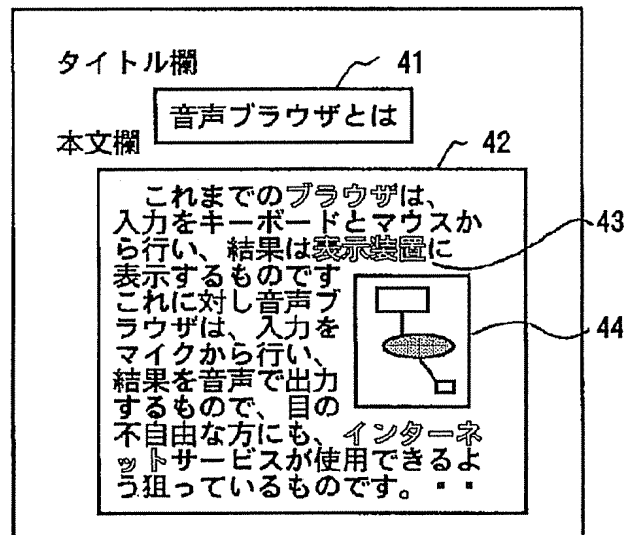
【図6】

本発明の第1の実施例の  
サーバクライアントシステムの動作シーケンスチャート

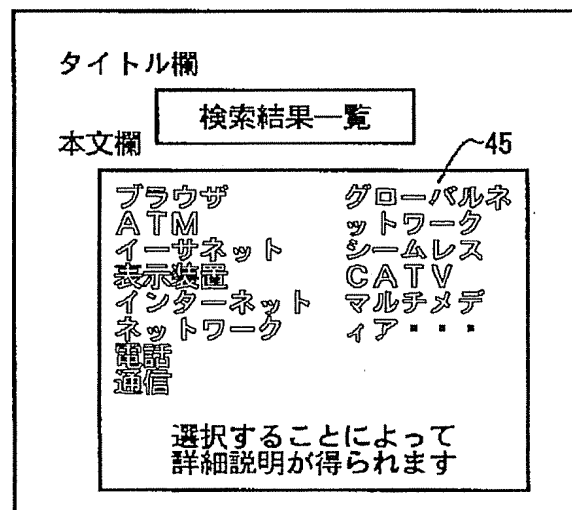


【図7】

## 本発明の第2の実施例における情報の型の例の説明図



(A) 混在型の場合

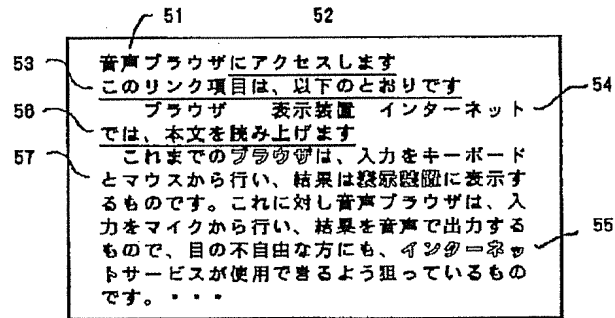


(B) 列挙型の場合

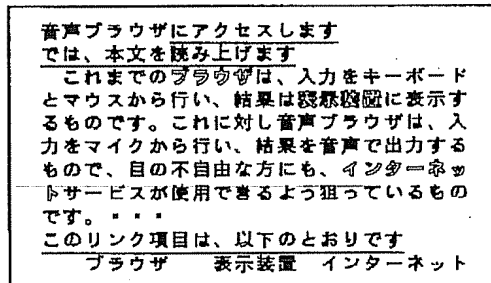


【図 8】

## 本発明の第 2 の実施例における混在型情報の読み上げ順の説明図

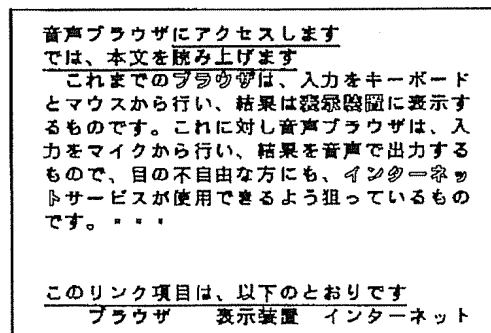


A) タイトル—リンク—本文順読み上げ方式



(B) タイトル—本文—リンク順読み上げ方式

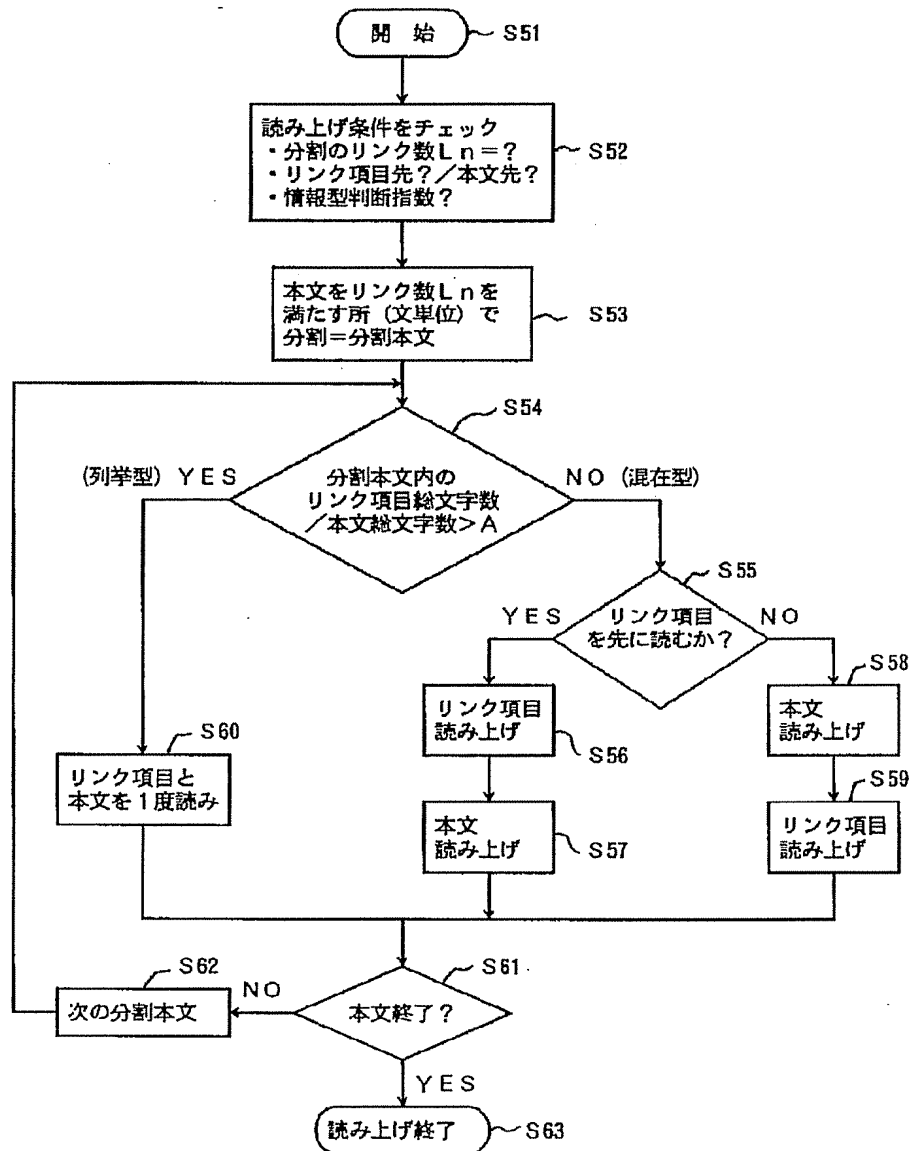
【リンク項目】  
と音声入力



(C) リンク項目読み上げ要求方式

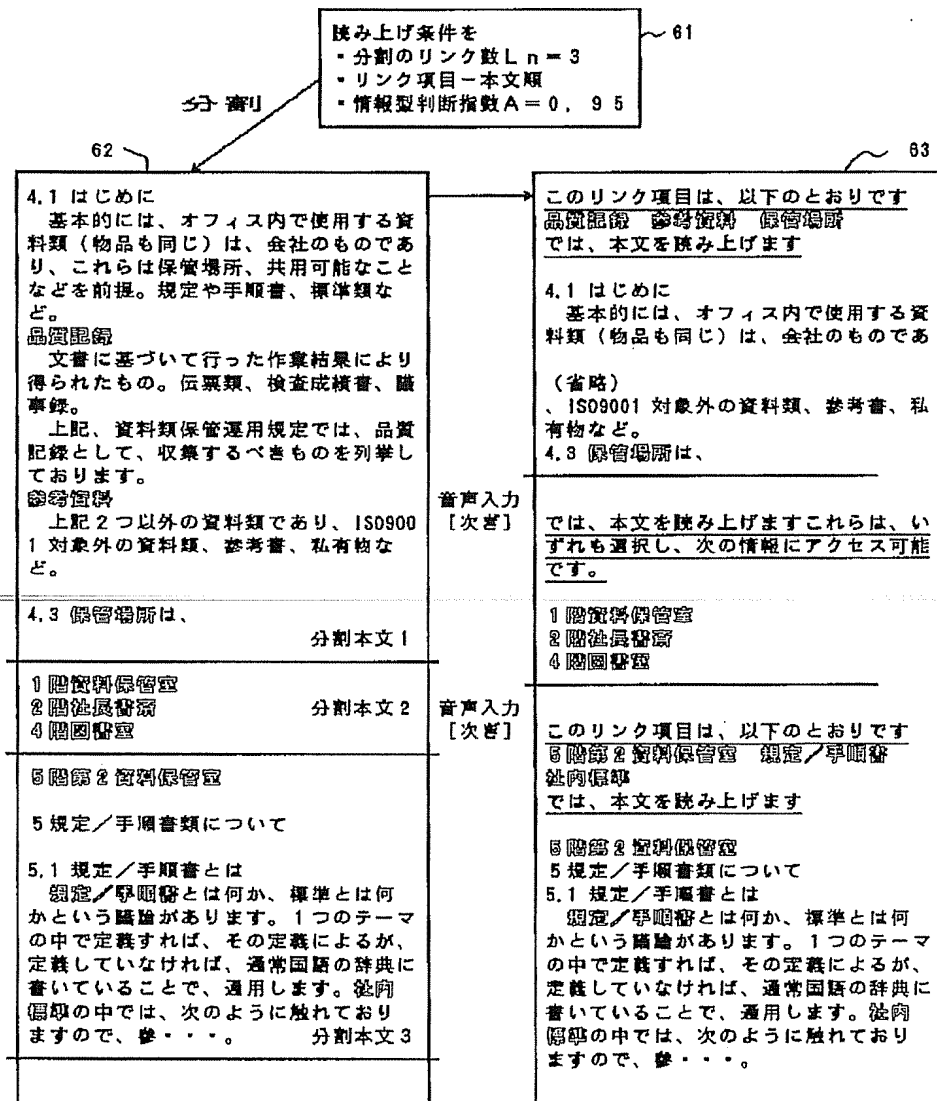
【図 10】

## 本発明の第 3 の実施例による読み上げ条件自動処理のフローチャート



【図11】

## 本発明の第4の実施例による分割本文と読み上げ内容との説明図



フロントページの続き

(51) Int. Cl. 6

識別記号

F I

G 0 6 F 15/60

(72) 発明者 木目田 常治

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72) 発明者 並木 育夫

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72) 発明者 氏家 正美  
東京都渋谷区桜丘町20番 1 号 エヌティテ  
ィエレクトロニクス株式会社内

(72) 発明者 虫壁 一彦  
東京都渋谷区桜丘町20番 1 号 エヌティテ  
ィエレクトロニクス株式会社内